

## Data Wrangling

The process of gathering, selecting, and programmatically transforming data into a format that makes it easier to work with. Data wrangling seeks to remove the risk of incomplete, unreliable & faulty data by ensuring that the data is in a reliable state before it's analyzed and leveraged. It enables businesses to tackle more complex data in less time, produce more accurate results, and make better decisions.



### Steps

◆ Discovering



◆ Structuring



◆ Cleaning



◆ Enriching



◆ Validating



◆ Publishing



### Use Cases

◆ Detecting Corporate Fraud

◆ Supporting Data Security

◆ Ensuring Accurate & Recurring Data Modeling Results

◆ Ensuring Business Compliance With Industry Standards

◆ Performing Customer Behaviour Analysis

◆ Reducing Time Spent On Preparing Data For Analysis

◆ Finding Out Data Trends

◆ Recognizing Business Value Of Data

### Trends

◆ The global data wrangling market size was valued at \$1,458.69 million in 2018 & is projected to reach \$5,581.18 million by 2026, growing at a CAGR of 18.40%

◆ Data wrangling costs analytics professionals as much as 80% of their time, leaving only 20% for exploration and modeling

◆ According to Gartner, 35% of large organizations will either buy or sell data via formal marketplaces by 2022, up from 25% in 2020